Problem 1: Let n be a positive integer. Use the Cauchy-Schwartz inequality to show that there are no real numbers $a_1, a_2, \ldots, a_n$ such that $\sum_{i=1}^{n} a_i^k = k$ for all positive integers $k > 1$.

Proof:

We will proceed by contradiction. Suppose that the set of real numbers $a_1, a_2, \ldots, a_n$ posesses the property that

$$\sum_{i=1}^{n} a_i^k = k$$

For all integers $k > 1$. Set $\vec{A} = (a_1^p, a_2^p, \ldots, a_n^p)$ and $\mathbb{I} = (1, 1, \ldots, 1)$. Then by the Cauchy Schwartz inequality:

$$(\vec{A} \cdot \mathbb{I})^2 \leq (\vec{A} \cdot \vec{A})(\mathbb{I} \cdot \mathbb{I})$$

$$\Rightarrow \left(\sum_{i=1}^{n} a_i^p\right)^2 \leq n \sum_{i=1}^{n} a_i^{2p}$$

$$\Rightarrow p^2 \leq 2np$$

Taking $p = 3n$:

$$9n^2 \leq 6n^2$$

Since $n \geq 1$:

$$\Rightarrow 9 \leq 6 \quad ↯$$

Thus, there can exist no set of real numbers $a_1, a_2, \ldots, a_n$ such that $\sum_{i=1}^{n} a_i^k = k$ for all $k > 1$. ∎

Problem 2: Let $H_0: \Sigma_1 = \Sigma_2 = \dots = \Sigma_g$, $H_a: \Sigma_i \neq \Sigma_j$ for some i, j. If $\Lambda = \max_{\theta \in \Theta_0} L(\theta) / \max_{\theta \in \Theta} L(\theta)$, show that

$$\Lambda = \prod_{l=1}^{g} \left(\frac{|S_l|}{|S_{pooled}|} \right)^{(n_l - 1)/2}$$

Proof:

We will proceed with a direct computation of Λ. First we compute $\max_{\theta \in \Theta_0} L(\theta)$. Under the null hypothesis this is equivalent to maximizing $L(\mu_1, \mu_2, \dots, \mu_g, \Sigma)$.

The Likliehood $L(\mu_1, \mu_2, \dots, \mu_g, \Sigma)$ is given by the product of the marginal density functions for $N_i(\mu_i, \Sigma)$ when samples $\vec{X}_i$ are known and μ_i, Σ are parameters. Thus we seek to maximize:

$$L(\mu_1, \dots, \mu_g, \Sigma) = \prod_{i=1}^{g} L_i(\mu_i, \Sigma)$$

$$= \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma|^{n_i/2}} \exp\left(-tr\left[\Sigma^{-1}\left(\sum_{j=1}^{n}(X_{ij} - \bar{X}_{ij})(X_{ij} - \bar{X}_{ij})^T + n_i(\bar{X}_i - \vec{\mu}_i)(\bar{X}_i - \vec{\mu}_i)^T\right)\right]/2\right)$$

Which is maximized for $\mu_i = \bar{X}_i$, thus:

$$\leq \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma|^{n_i/2}} \cdot \exp\left(-tr\left[\Sigma^{-1} \sum_{j=1}^{n} (X_{ij} - \bar{X}_{ij})(X_{ij} - \bar{X}_{ij})^T\right]/2\right)$$

$$= \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma|^{n_i/2}} \exp\left(-tr\left[\Sigma^{-1}(n_i - 1) S_i\right]/2\right)$$

$$= \left(\prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma|^{n_i/2}}\right) \prod_{i=1}^{g} \exp\left(-tr\left[\Sigma^{-1}(n_i - 1) S_i\right]/2\right)$$

$$= \left(\prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma|^{n_i/2}} \right) \exp\left(-tr\left[\Sigma^{-1} \sum_{i=1}^{g} (n_i - 1) S_i\right]/2\right)$$

$$= \left(\prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma|^{n_i/2}} \right) \exp\left(-tr\left[\Sigma^{-1} S_{pool} \sum_{i=1}^{g} (n_i - 1)\right]/2\right)$$

$$= \left(\prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma|^{n_i/2}} \right) \left(\prod_{i=1}^{g} \exp\left(-tr\left[\Sigma^{-1} S_{pool} (n_i - 1)\right]/2\right)\right)$$

$$= \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma|^{n_i/2}} \exp\left(-tr\left[\Sigma^{-1} (n_i - 1) S_{pool}\right]\right)$$

Which, by the maximization result 4.10, has maximum value:

$$\prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |(n_i - 1) S_{pool}|^{n_i/2}} (n_i)^{p n_i/2} e^{-p n_i/2}$$

Thus:

$$\max_{\theta \in \Theta_0} L(\theta) = \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |(n_i - 1) S_{pool}|^{n_i/2}} (n_i)^{p n_i/2} e^{-p n_i/2}$$

Next we compute the Likelihood $\max_{\theta \in \Theta} L(\theta)$. This is given by:

$$\max_{\theta \in \Theta_0} L(\theta) = \max_{\mu_i, \Sigma_i} \prod_{i=1}^{g} L(\mu_i, \Sigma_i)$$

$$= \max_{\mu_i, \Sigma_i} \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma_i|^{n_i/2}} \exp\left(-tr\left[\Sigma_i^{-1} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T + n(\bar{x}_i - \mu_i)(\bar{x}_i - \mu_i)^T\right]/2\right)$$

$$= \max_{\Sigma_i} \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma_i|^{n_i/2}} \exp\left(-tr\left[\Sigma_i^{-1} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T\right]/2\right)$$

$$= \max_{\Sigma_i} \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |\Sigma_i|^{n_i/2}} \exp\left(-tr\left[\Sigma_i^{-1} (n_i-1) S_i\right]\right)$$

Which by the maximization result 4.10, is

$$= \prod_{i=1}^{g} \frac{1}{(2\pi)^{n_i p/2} |(n_i-1) S_i|^{n_i/2}} (n_i)^{pn_i/2} e^{-pn_i/2}$$

Thus, the ratio becomes:

$$\Lambda = \max_{\theta \in \Theta_0} L(\theta) / \max_{\theta \in \Theta} L(\theta)$$

$$= \prod_{i=1}^{g} \frac{(2\pi)^{n_i p/2} |(n_i-1) S_i|^{n_i/2}}{(n_i)^{pn_i/2} e^{-pn_i/2}} \cdot \frac{(n_i)^{pn_i/2} e^{-pn_i/2}}{(2\pi)^{n_i p/2} |(n_i-1) S_{pool}|^{n_i/2}}$$

$$= \prod_{i=1}^{g} \left(\frac{|S_i|}{|S_{pool}|} \right)^{n_i/2}.$$

Problem 3: Let $Y \sim N_n(Z\beta, \sigma^2 I)$ be a linear regression model where Z is a $n \times (p+1)$ matrix of full rank. Test the hypothesis $H_0: \beta_1 + \beta_2 - 2\beta_3 = 0$ vs. $H_1: \beta_1 + \beta_2 - 2\beta_3 \neq 0$.

Let the matrix C be defined as

$$C = [1, 1, -2, 0, 0, \ldots, 0]_{1 \times (p+1)}$$

Then, our null hypothesis is $C\beta = 0$, where the number of linear combinations tested is $r - q = 1$.

By Result 7.6, we can test this hypothesis via the ratio test (p. 376)

$$\frac{(C\hat{\beta})^T (C(Z^T Z)^{-1} C^T)^{-1} (C\hat{\beta})}{n\sigma^2/(n-r-1)} > F_{1, n-r-1}(0.05)$$

Where the above inequality holding would imply we reject the null hypothesis with an $\alpha = 0.05$.

Note that since C is of dimension $1 \times (p+1)$, $\hat{\beta}$ of dimension $(p+1) \times 1$, and $(Z^T Z)^{-1}$ of dimension $(p+1) \times (p+1)$, all terms contained in parenthesis of the inequality are scalars. Thus:

$$\frac{(C\hat{\beta})^T (C(Z^TZ)^{-1}C^T)^{-1} (C\hat{\beta})}{n\sigma^2/(n-r-1)} > F_{1,n-r-1}(0.05)$$

$$\Rightarrow (C\hat{\beta})^T (C(Z^TZ)^{-1}C^T)^{-1} (C\hat{\beta}) > \frac{n\sigma^2 F_{1,n-r-1}(0.05)}{n-r-1}$$

$$\Rightarrow (C\hat{\beta})^T (C\hat{\beta}) > \frac{n\sigma^2 F_{1,n-r-1}(0.05)}{n-r-1} [C(Z^TZ)^{-1}C^T]$$

$$\Rightarrow (\hat{\beta}_1 + \hat{\beta}_2 - 2\hat{\beta}_3)^2 > \frac{n\sigma^2 F_{1,n-r-1}(0.05)}{n-r-1} [C(Z^TZ)^{-1}C^T]$$

And we reject the null hypothesis $\hat{\beta}_1 + \hat{\beta}_2 - 2\hat{\beta}_3 = 0$ if the above inequality holds.

Problem 5: Find the smallest and largest values of the expression:

$$x^T A x / x^T x \quad \text{where } A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \text{ and } x \neq 0.$$

By the maximization of Quadratic Forms for Points on the Unit sphere (p. 118), we know that $x^T A x / x^T x$ has maximum λ_1 and minimum λ_3, where $\lambda_1 > \lambda_2 > \lambda_3$ are the eigenvalues of A.

Thus, we search for the eigenvalues:

$$|A - \lambda I| = 0$$

$$\Rightarrow \left|\begin{array}{ccc|cc} 2-\lambda & -1 & 0 & 2-\lambda & -1 \\ -1 & 2-\lambda & -1 & -1 & 2-\lambda \\ 0 & -1 & 2-\lambda & 0 & -1 \end{array}\right| = (2-\lambda)^3 - [2-\lambda + 2-\lambda] = 0$$

$$\Rightarrow (2-\lambda)^3 - 4 + 2\lambda = 0$$

$$\Rightarrow (2-\lambda)^3 + 2(-2+\lambda) = 0$$

$$\Rightarrow (2-\lambda)^3 - 2(2-\lambda) = 0$$

$$\Rightarrow (2-\lambda)\left[(2-\lambda)^2 - 2\right] = 0$$

$$\Rightarrow \lambda = 2 \quad \text{or} \quad 4 - 4\lambda + \lambda^2 - 2 = 0$$

$$\Rightarrow \lambda^2 - 4\lambda + 2 = 0$$

$$\Rightarrow \lambda = \frac{4 \pm \sqrt{16-8}}{2}$$

$$\Rightarrow \lambda = 2 \pm \frac{2\sqrt{2}}{2}$$

$$\Rightarrow \lambda = 2 \pm \sqrt{2}$$

Since $2+\sqrt{2} > 2 > 2-\sqrt{2}$, we have that the maximum of $x^T A x / x^T x$ is $2+\sqrt{2}$, and the minimum is $2-\sqrt{2}$.

Problem 6: Excercise 6.11 on pg. 339

Proof:

We need to find the values for $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\Sigma}$ which maximize the likelihood $L(\mu_1, \mu_2, \Sigma)$.

This likelihood is given by the product of the marginal densities $N_p(\mu_1, \Sigma)$ and $N_p(\mu_2, \Sigma)$, when samples X_1 and X_2 are known and μ_1, μ_2, Σ are parameters. This is (4-16) on P. 170 and is:

$$L(\mu_1, \mu_2, \Sigma) = \prod_{i=1}^{2} L(\mu_i, \Sigma)$$

$$= \prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} |\Sigma|^{n_i/2}} e^{-tr\left[\Sigma^{-1} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T + n(\bar{x}_i - \vec{\mu}_i)(\bar{x}_i - \vec{\mu}_i)^T\right]/2}$$

This function is maximal for the smallest possible value in the exponent. Since we are allowed to vary $\vec{\mu}_i$, this minimum value in the exponent is obtained at $\vec{\mu}_i = \bar{x}_i$, since this would make the strictly non-negative second term of $\sum_{i=1}^{n} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T + n(\bar{x}_i - \vec{\mu}_i)(\bar{x}_i - \vec{\mu}_i)^T$ equal to 0.

Thus:

$$\max L(\vec{\mu}_1, \vec{\mu}_2, \Sigma) = \max \prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} |\Sigma|^{n_i/2}} e^{-tr\left[\Sigma^{-1} \sum_{i=1}^{n} (x_{ij} - \bar{x})(x_{ij} - \bar{x})^T\right]/2}$$

$$= \max_{\Sigma} \prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} |\Sigma|^{n_i/2}} e^{-tr\left[\Sigma^{-1} (n_i - 1) S_i\right]/2}$$

$$= \max_{\Sigma} \left[\left(\prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} \, |\Sigma|^{n_i/2}} \right) \prod_{i=1}^{2} e^{-tr[\Sigma^{-1}(n_i - 1) S_i]/2} \right]$$

$$= \max_{\Sigma} \left[\left(\prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} \, |\Sigma|^{n_i/2}} \right) e^{-tr[\Sigma^{-1}(n_1 - 1) S_1 + \Sigma^{-1}(n_2 - 1) S_2]/2} \right]$$

$$= \max_{\Sigma} \left[\left(\prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} \, |\Sigma|^{n_i/2}} \right) e^{-tr[\Sigma^{-1}\{(n_1 - 1) S_1 + (n_2 - 1) S_2\}]/2} \right]$$

$$= \max_{\Sigma} \left[\left(\prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} \, |\Sigma|^{n_i/2}} \right) e^{-tr[\Sigma^{-1} S_{pooled} \sum_{j=1}^{2} (n_j - 1)]/2} \right]$$

$$= \max_{\Sigma} \left[\left(\prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} \, |\Sigma|^{n_i/2}} \right) \prod_{i=1}^{2} e^{-tr[\Sigma^{-1} S_{pooled} (n_i - 1)]/2} \right]$$

$$= \max_{\Sigma} \prod_{i=1}^{2} \frac{1}{(\pi)^{n_i p/2} \, |\Sigma|^{n_i/2}} e^{-tr[\Sigma^{-1} S_{pooled} (n_i - 1)]/2}$$

$$= \max_{\Sigma} \frac{1}{(\pi)^{\frac{n_1 p + n_2 p}{2}} |\Sigma|^{(n_1 + n_2)/2}} e^{-tr[\Sigma^{-1} S_{pooled} (n_1 - 1) + \Sigma^{-1} S_{pooled} (n_2 - 1)]/2}$$

$$= \max_{\Sigma} \frac{1}{(\pi)^{(n_1 p + n_2 p)/2} \, |\Sigma|^{(n_1 + n_2)/2}} e^{-tr[\Sigma^{-1} S_{pooled} (n_1 + n_2 - 2)]/2}$$

Which by the maximization result 4.10, achieves its maximum for $\hat{\Sigma} = \frac{1}{n_1 + n_2} S_{pooled} (n_1 + n_2 - 2) = \left(\frac{n_1 + n_2 - 2}{n_1 + n_2} \right) S_{pooled}$.

Along with our previously mentioned maximizing values of $\vec{\mu}_1$, $\vec{\mu}_2$, we have that we achieve a maximum likelihood for:

$$\hat{\mu}_1 = \bar{X}_1 \ , \ \hat{\mu}_2 = \bar{X}_2 , \ \hat{\Sigma} = \left(\frac{n_1 + n_2 - 2}{n_1 + n_2} \right) S_{pooled}$$